Title

Robust cloud motion estimation by spatio-temporal correlation analysis of irradiance data

Authors

Jamaly, Mohammad
Kleissl, Jan

Publication Date

2018

DOI

10.1016/j.solener.2017.10.075

Peer reviewed

Robust Cloud Motion Estimation by Spatio-temporal Correlation Analysis of Irradiance Data

Mohammad Jamaly and Jan Kleissl

Center for Renewable Resources and Integration, Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA 92093-0411, United States

Abstract

The main contributor to spatio-temporal variability in the solar resource is clouds passing over photovoltaic (PV) modules. Cloud velocity is a principal input to many short-term forecast and variability models. In this paper spatio-temporal correlations of irradiance data are analyzed to estimate cloud motion. The analysis is performed using two spatially and temporally resolved simulated irradiance datasets generated from large eddy simulation. Cloud motion is estimated using two different methods; the cross-correlation method (CCM) applied to two or a few consecutive time steps and cross-spectral analysis (CSA) where the cloud speed and direction are estimated by cross-spectral analysis of a longer time series. CSA is modified to estimate the cloud motion direction as the case with least variation for all the velocities in the cloud motion direction. To ensure reliable cloud motion estimation, quality control (QC) is added to the CSA and CCM analyses. The results show 33% (52°) and 21% (6°) improvement in the cloud motion speed (direction) estimation using the modified CSA and CCM over the original methods (without QC), respectively. In general, CCM results are accurate for all the different cloud cover fractions with average relative mean bias error (rMBE) of cloud speed and mean absolute error of cloud direction equal to 3% and 3°, respectively. For low cloud cover fractions, CSA estimates the cloud motion speed and direction with rMBE and mean absolute error equal to 10% and 11°, respectively. However, for high cloud cover fractions and unsteady cloud speed, CSA results are not reliable for 3-4 h time series; however, splitting the whole time series into shorter time intervals reduces the rMBE and mean absolute error to 15% and 16° respectively.

Keywords: Solar radiation; Solar forecast; Spatio-temporal variability; Cloud motion.

1. Introduction

1.1. Motivation

The power output from solar photovoltaic (PV) power plants is usually more variable than conventional power generation sources. Variability is the main challenge for integration of large amounts of PV power plants into the electricity grid (Marcos et al., 2011). The ability to forecast

actual variability of solar distributed generation (DG) will allow grid operators to better accommodate the variable electricity generation for resource adequacy considerations, such as scheduling and dispatching of power.

Besides predictable solar variability according to diurnal and annual irradiance patterns, the main source of spatio-temporal variability in the solar resource is transient clouds and that variability is related to the cloud optical depth and speed. Cloud motion is the main input to most short-term solar variability and forecast models (Arias-Castro et al., 2014; Hoff and Perez, 2010; Lave and Kleissl, 2013; Chow et al., 2011; Marquez and Coimbra, 2013; Perez et al., 2010; Yang et al., 2014a; Yang et al., 2014b; Lorenzo et al., 2014). Therefore, cloud motion estimation has been extensively investigated recently (Bosch et al., 2013; Bosch and Kleissl, 2013; Fung et al., 2013; Huang et al., 2013; Quesada-Ruiz et al., 2014; Chow et al., 2015). Accurate cloud motion vectors are critical for solar forecast, interpolation, and variability analyses (Jamaly and Kleissl, 2017).

1.2. Cloud motion estimation

Vega-Riveros and Jabbour (1989) reviewed various techniques related to the motion analysis and detection. Motion analysis methods are either based on the direct numerical solution of the optical flow constraint equation (method of differentials) or correspondence-based approaches, where image features are identified and tracked to measure their displacement. These measurements are then used to calculate the displacement of the object as a whole. Estimating cloud motion for sky imaging and satellite data by solving the optical flow equation incurs less computational expense. However, it has many restrictions. Therefore, most of the methods for estimation of the Cloud Motion Vectors (CMVs) are developed using correspondence-based approaches. In general, CMVs are obtained by first locating salient image features such as brightness gradients, corners, cloud edges, or brightness temperature gradients (Bedka and Mecikalski, 2005; Menzel, 2001). Then, assuming the features do not change significantly over a short interval, CMVs are computed by tracking the features in successive images.

CMVs have been obtained using sky imaging devices (Marquez et al., 2013) for very short-term solar forecasts up to 20 min ahead. Moreover, CMVs have been estimated from satellite imagery (Perez and Hoff, 2013; Menzel, 2001; Hammer et al., 1999; Leese et al., 1971). Escrig et al. (2013) applied multispectral tests and binary cross-correlations for cloud motion estimation using geostationary satellite imagery. They applied coherence and quality control tests to the resulting motion vectors and proposed new thresholds for infrared and visible tests. Fuh and Maragos (1991) developed a model for estimating the displacement field in spatio-temporal image sequences that allows for affine cloud shape deformations. The model is based on the block matching method (which is based on the same principal as the cross-correlation method presented later) and parameters were found using a least-squares algorithm. Post-smoothing the velocity field via spatio-temporal vector median filtering almost always improves the performance of the matching algorithm. However, block matching has a higher computational complexity.

Farnebäck (2003) developed a method for motion estimation based on a two-frame algorithm. The first step is to approximate each neighborhood of both frames by quadratic polynomials. Then, a method to estimate displacement fields from the polynomial expansion coefficients was derived. The main weakness of the algorithm is the assumption of a slowly varying displacement field, causing discontinuities to be smoothed out. Hammer et al. (1999) developed a statistical method based on conditional probabilities to compute CMVs and predict solar radiation up to 2 h ahead. Lorenz et al. (2004) used a similar method (applying extrapolation of motion assuming

persistence of cloud speed, size, and shape) to obtain solar radiation forecasts up to 6 hours ahead. For longer forecast time horizons, non-linearities in atmospheric motion and cloud formation and evaporation cause Numerical Weather Prediction (NWP) models to outperform satellite-based CMV forecasts (Perez et al., 2010). Arking et al. (1978) applied Fourier phase difference technique which allows motion estimates to be made for individual spatial frequencies related to cloud pattern dimensions. However in the presence of mixtures of motions, changes in cloud shape and edge effects, the cross-correlation scheme yields a more reliable estimate of cloud motion than the phase difference technique.

Since CMV estimation by either sky imaging, satellite data, or NWP lack granularity and computational efficiency, local ground measurements of cloud speed are advantageous for short-term solar variability and solar forecasting (Bosch et al., 2013). Bosch and Kleissl (2013) showed that cloud motion can be detected from spatio-temporal irradiance or power measurements across a utility-scale PV plant from the timing of cloud arrival at three different points.

1.3. The proposed method

Prior methods using ground data were predicated upon sparse data. The analysis in this paper is motivated by the increased availability of dense PV power output observations in urban areas with spatial resolution on the order of 100s of meters. Actual PV power output can be converted to clear sky index (see e.g. Engerer and Mills, 2014) and then cloud motion could be estimated just like if the PV system was an irradiance sensor. Therefore the success of two algorithms in detecting cloud motion is estimated from simulated dense ground data: cross-spectral analysis (CSA) and the cross-correlation method (CCM). In CSA, the cloud speed and direction are estimated by cross-spectral analysis of the irradiance data at some given locations (sites) through the domain (Inoue et al., 2012; Shinozaki et al., 2014). The CSA method suggested by Inoue et al. (2012) and Shinozaki et al. (2014) is restricted by the spatial arrangement of the sites such that the cloud direction may be inaccurate if there are only a few distinct relative angles between the pairs of the chosen sites. To remove the restriction, a new CSA approach for cloud motion direction is proposed by selecting the direction with least variation for all the velocities in the cloud motion direction.

In CCM, the cloud motion is estimated by comparing correlation between spatial irradiance data at two or more time steps (Hamill and Nehrkorn, 1993). The CCM suggested by Hamill and Nehrkorn (1993) is generalized for cloud movement estimation using unstructured ground measured data. Moreover, to compare the consistency of the method when applied to different scales, CCM is applied by considering the whole domain as well as smaller subdomains. Also, to ensure reliable cloud motion estimation, quality control (QC) is added to the CSA and CCM analyses including removing conditions with low variability and less correlated sites.

The algorithms are tested only on simulated ground data, which is advantageous because the true cloud speed is known. In real datasets the true cloud speed is unknown and such data suffer from spatial heterogeneity in surface and atmospheric conditions that manifests in spatial differences in cloud motion vectors. Such heterogeneities can be avoided in a simulated dataset and the cloud motion estimation results are therefore expected to be more generalizable. In Jamaly and Kleissl (2017) the CSA and CCM methods are applied to real data for spatio-temporal interpolation or forecast of solar irradiance.

The datasets are described in section 2. The cloud speed methodology is described in section 3. Results of the estimation of the cloud motion are presented in section 4 and section 5 concludes the paper.

2. Dataset

The analysis has been performed using two spatially and temporally resolved simulated irradiance datasets generated from large eddy simulation (LES). LES is a three-dimensional computational fluid mechanics technique that numerically integrates the Navier-Stokes equations. The momentum, temperature, and moisture transport is simulated at each grid point. High spatial and temporal resolution allows simulating the large turbulent motions in the atmospheric boundary layer explicitly and LES therefore produces more accurate wind, temperature, moisture, and cloud fields than other techniques. Periodic boundary conditions in the horizontal directions are used to represent an infinitely long, homogeneous domain that allows atmospheric turbulence to develop in a realistic manner. LES is forced by a geostrophic wind at the top of the domain. Surface fluxes of heat and water largely determine the relative humidity in the boundary layer and whether clouds will form. We apply the well-validated UCLA-LES using the same settings as Ghonima et al. (2016). Simulated datasets are considered since LES wind vectors at the average cloud height can be considered as the reference cloud motion.

2.1. RICO simulation

In the first simulation, a spatial domain of 2540 m x 2540 m (128 x 128 grid points) with boundary and initial conditions from the rain in cumulus over the ocean (RICO) field study (vanZanten et al., 2011) centered at 18.0° N, 61.8° W is setup. The simulation is performed up to 4000 m height resolved by 100 grid points. The precipitating RICO case with boundary layer moisture in the initial profile equal to 12.35 g/kg is simulated. Following 4 hours of spinup, 10 sec liquid water path (LWP) aggregated from cloud base to cloud top is output over a 30 min interval. Also, a representative wind speed vector is output at each time step; the two velocity components are $u(x,y,z_c,t)$ and $v(x,y,z_c,t)$, where z_c is average cloud height. The wind velocity is considered as the reference cloud motion and compared against estimated cloud motion in Section 4.

2.2. CGILS simulation

The second simulated dataset consists of simulated LWP obtained by the CGILS (CFMIP-GCSS Intercomparison of Large-Eddy and Single-Column Models) stratocumulus cloud over-land case (see Zhang et al., 2009) with Bowen ratio of 0.1 (moist surface) near the California Coast (35° N, 125° W). The spatial domain covers 2400 m x 2400 m resolved by 96 x 96 grid points in the horizontal and extends to 965 m height in the vertical using 193 grid points. 10 sec resolution LWP is output for 24 hours (8641 time steps) starting at midnight. Also, a representative wind speed vector at the average cloud height is output every 1200 sec. Since the cloud deck is overcast till noon and breaks up in the afternoon, the cloud speed is estimated separately during 0800-1200 h and 1200 - 1500 h. The CGILS dataset presents more challenging conditions for cloud motion estimation than RICO: (i) Overcast stratocumulus (CGILS) lack the sharp transition between cloud boundaries and clear sky that are common for cumulus; (ii) Cloud motion is unsteady for CGILS while it is steady for RICO. Unsteady cloud motion is a challenge for the CSA method, which relies on temporal averaging.

2.3. Converting LWP to clear sky index

The 1D shortwave radiation scheme proposed by Slingo (1989) is applied to calculate total transmission coefficient (equivalent to clear sky index *Kt*, the ratio of global irradiation to the global irradiation in clear sky condition) using the LWP and the equivalent radius of the drop

size distribution (r_e). According to Hess et al. (1998), the typical effective radius for cumulus clouds is 5.8 μm (clean) and 4.0 μm (polluted), so 5 μm is chosen for the first simulation. The effective radius was set to be 12.7 μm for the stratocumulus cloud in the second simulation. Note that the Slingo model with 4 spectral bands is applied for simplicity but the full 24 bands model can be used as well. The inverse relation between LWP and *Kt* was confirmed visually.

3. Cloud speed methodology

3.1. Cross-spectral analysis (CSA)

In CSA, the cloud speed and direction are estimated by analyzing the time series of the irradiance data at the given sites through the domain (Inoue et al., 2012 and Shinozaki et al., 2014).

3.1.1. Classical CSA definition

The cloud speed and direction are estimated by CSA of the observed *Kt* at some given locations. As suggested by Inoue et al. (2012) and Shinozaki et al. (2014), for each pair of sites *i* with $i = 1, \ldots, n$, the time difference ΔT_i that maximizes the delayed (cross) correlation between the two *Kt* time series is obtained. Then, the time delay moment $M(\theta)$ for all the sites with angle θ is calculated using the delay times for all pairs of sites as

$$M(\theta) = \sum_{\theta-30<\theta_i<\theta+30} \Delta T_i / r_i, \qquad \text{Eq. (1)}$$

where r_i is the distance between the i^{th} pair of sites with angle θ_i. The movement angle θ_m is calculated such that the moment $M(\theta)$ of the time delay is maximized. After that, for each pair of sites, the velocity vector in the θ_m direction (V_i) is calculated as

$$V_i = \frac{r_i \cos(\theta_i - \theta_m)}{\Delta T_i}, \qquad \text{Eq. (2)}$$

and the cloud speed is chosen as the median of the V_i of all pairs of sites.

However, there may be no local maximum in the moment and then it is difficult to estimate the movement speed and direction. If there is no clear pattern in the delay time, the moment would approach 0. As suggested by Shinozaki et al. (2014), an error index *e* quantifies the confidence

$$e = \frac{V(60\%) - V(40\%)}{V(50\%)} \times 100, \qquad \text{Eq. (3)}$$

where $V(50\%)$ is the median of V_i and $V(40\%)$ and $V(60\%)$ are the 40% and 60% percentiles of V_i respectively. Shinozaki et al. (2014) mentioned that when a clear movement trend is observed (there is a dominant cloud motion direction), the error index *e* was usually less than 40%. An unclear movement trend (there are some different cloud directions with almost the same values of the time delay moment) was correlated with error index *e* of about 40% to 70%; when there was no movement trend (cloud motion directions are too different), the error index *e* was greater than 70%.

3.1.2. Robust estimation of the cloud motion

The approach suggested by Inoue et al. (2012) and Shinozaki et al. (2014) is restricted by the spatial arrangement of the sites. This is because the movement angle θ_m is the average angle between all sites. In the worst case scenario all sites are in a line with angles equal to 0 or 180, so the moment has nonzero values only at these two points which is not necessarily the cloud motion direction. Therefore the number and range of possible values of θ_m are limited by the diversity of angles between the sites.

Instead of the time delay moment in the original CSA (Eq. 1), we propose a new approach for cloud motion direction. For a given movement angle θ_s, the velocity detected by each pair of sites is given by Eq. 2. Ideally, the velocity components in the given movement direction for all pairs of the sites should be equal when $\theta_s = \theta_m$. So, the best candidate for movement direction is the case with least variation for all the velocities. For a given θ_s, the variation for all the velocities is calculated as

$$V_{var}(\theta_s) = var\left(\frac{r_i \cos(\theta_i - \theta_s)}{\Delta T_i}\right), \qquad \text{Eq. (4)}$$

where the sample variance is calculated as

$$var = \frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \bar{Z})^2, \qquad \text{Eq. (5)}$$

in which $Z_i = [r_i \cos(\theta_i - \theta_s)]/\Delta T_i$ and $\bar{Z} = [\sum_{i=1}^{n} Z_i]/n$.

The time delay variation moment $V_{var}(\theta_s)$ is calculated for all angles (θ = 1°, …, 360°) and the angle for which the minimum occurs is the cloud movement direction (θ_m). Also, the cloud movement speed is considered as the median of velocities at this direction (Eq. 2).

To get a realistic time delay ΔT_i for each pair of sites (which is applied in Eq. 4), a maximum limit for allowable time lag ($\max_{lag}$) is needed. $\max_{lag}$ is critical since a small $\max_{lag}$ could be shorter than the physical cloud transit time while a large $\max_{lag}$ causes the overlapping length of the lagged time series to become too small and, therefore, the obtained ΔT_i is not reliable. Since irradiance signals from clear to cloudy (and vice versa) are similar there is a risk of spurious correlation peaks from two different clouds passing two sites. Therefore, different maximum allowable time lags are applied to achieve robust results; $\max_{lag}$ = [1, 2, 3, 4, 5, 6] t_{dim}/6, where t_{dim} is length of time series. Cloud movement direction, speed and error index (Eqs. 4, 2 and 3 respectively) are obtained for each $\max_{lag}$ and the best results are considered as the case with the smallest error index. A problem that is unique to this simulated dataset with periodic boundary conditions is that a site is both upwind and downwind of other sites as the motion across the domain boundary is considered. Such correlation peaks associated with motion across domain boundaries are discarded.

3.1.3. Excluding sites and data

Quality control (QC) improves the robustness and accuracy of CSA as some pairs of sites have low quality data. Including these low quality data may lead to inaccurate and unreliable cloud motion estimation. Therefore, a new QC approach is developed and applied in the analysis as follows. These steps are critical for reliable results. The specific thresholds in the QC are all chosen empirically, but appear to yield good results for the wide range of cloud conditions simulated here.

A. Remove conditions with low variability

When the variability in *Kt* is small such as in clear conditions, the signal is insufficient to detect cloud direction and speed. Therefore, for each site with given *Kt* time series (Kt_s), the variation ratio of the site (var_s) is calculated as

$$var_s = 1 - \frac{\text{mean}(Kt_s)}{\max(Kt_s)}. \quad \text{Eq. (6)}$$

Also, for each site, the cloud cover fraction is calculated as the fraction of time steps with Kt_s < 0.85. It is empirically found that sites with low variation ratio (var_s < 0.1) or low cloud cover fraction (less than 0.1) lead to inaccurate cloud motion estimation and, hence, these sites are removed from the analysis. In the RICO simulation with total 128 x 128 = 16,384 sites (pixels), 11,103 sites (68%) with low variation ratio or low cloud cover fraction are removed from the analysis, while for the CGILS simulation 5,571 (60%) and 6,318 (69%) out of 9,216 sites are removed during cloud breakup and overcast periods respectively.

B. Remove site pairs that are too distant

The cross-correlation is applied on the remaining pairs of sites. Note that for a given maximum time lag, $\max_{lag}$, there is a limit on the cloud speed that can be detected between each pair of sites: the cloud speed for a given pair of sites with distance r_i can only take on the values of $r_i/\max_{lag}$, 2 $r_i/\max_{lag}$, … , and r_i. If the pair of sites is too distant and/or the length of time series (and therefore $\max_{lag}$) is relatively short ($r_i/\max_{lag} >> 1$), the velocity between the pair of sites may be overestimated. To ensure reliable velocity field estimation, a limit on distance between pairs of sites should be applied (d_{lim}). Without loss of generality and to be able to detect velocity values down to 1 m/s, the limit is applied such that $r_i/\max_{lag}$ = 1 m/s. Therefore, for a given $\max_{lag}$, d_{lim} is calculated as

$$d_{lim} = max_{lag} \times 1\frac{\text{m}}{\text{s}}. \quad \text{Eq. (7)}$$

To improve accuracy of the results and robustness of the method, site pairs with distance greater than d_{lim} are not considered in the CSA. For instance, for the case with $\max_{lag} = t_{dim}$, 5.1 x 10^6 (or 36%) of the remaining 1.4 x 10^7 pairs of sites, 2.4 x 10^6 (or 37%) of the remaining 6.6 x 10^6 pairs of sites, and 1.5 x 10^6 (or 37%) of the remaining 4.2 x 10^6 pairs of sites are removed from the analysis in RICO simulation, CGILS simulation during cloud breakup, and overcast periods respectively.

C. Remove less correlated sites

For each pair of sites, maximum and average values of the cross-correlation of the pair (max_{QC} and $mean_{QC}$) are calculated. The cross-correlation QC ratio (r_{QC}) is then

$$r_{QC} = 1 - \frac{mean_{QC}}{max_{QC}}. \quad \text{Eq. (8)}$$

max_{QC} shows the degree of similarity between the two time series. r_{QC} represents the level of the certainty of the cross-correlation between the site pair; large r_{QC} means that there is a large maximum cross-correlation relative to the average and, therefore, the cross-correlation result is more reliable. To ensure high quality results, the site pairs with max_{QC} < 0.8 or r_{QC} < 0.8 are removed from the analysis. These empirically-derived thresholds are found to be sufficient for accurate results in all cases. Therefore, for $\max_{lag} = t_{dim}$ case, 1.1 x 10^6 (out of the remaining 8.8

x 10^6 pairs of sites), 5.0 x 10^5 (out of the remaining 4.2 x 10^6 pairs of sites), and 3.2 x 10^5 (out of the remaining 2.7 x 10^6 pairs of sites) pairs of sites are removed from the analysis in RICO simulation, CGILS simulation during cloud breakup and overcast periods, respectively.

3.1.4. Reduce the density of sites

To reduce computational costs, it will be investigated whether choosing less points with the highest local variabilities are sufficient to achieve the same results. This capability of the method would be advantageous in practice since the site density is usually lower than in the simulated case. Therefore, the whole domain is split into boxes with 5 x 5 pixels and at each box the point with the highest variation ratio (as defined in Eq. 6) is chosen (625 and 361 sites for the RICO and CGILS simulations respectively). Geographically restricting the selection of points ensures that a diverse set of pairs with many distances and directions is available.

3.2. Cross-correlation method (CCM)

Hamill and Nehrkorn (1993) proposed calculating the cloud velocity and direction of motion through the cross-correlation method (CCM) applied to two consecutive time steps. Hamill and Nehrkorn (1993) applied CCM on satellite data to forecast cloud movement. Here, the method is generalized for cloud movement estimation using unstructured ground measured data by interpolating *Kt* data on a regular structured grid.

The CCM finds the position that best matches each given subset of pixels at the previous time step given the data at the current time step. The CCM yields a cloud vector (direction and speed) with the largest cross-correlation coefficient (CCC) that quantifies the quality of the match. Besides CCC, there are other cost functions including sum of the absolute value of difference (SAVD) and sum of squared difference (SSD); however, CCC is the most robust (Yang et al., 2012). In general, as mentioned by Chow et al. (2011), the CCM obtains accurate motion in heterogeneous areas with high contrast of pixel values, such as clouds with a sharp boundary in a clear sky. Quality control is essential to detect such cases as described below.

3.2.1. Calculating velocity components in different domains

The cloud velocity is estimated using data of the whole domain (X_{dim} and Y_{dim} are dimension in X and Y directions, respectively) and smaller rectangular subsets (boxes). The box sizes and searching radii (X_{radius} and Y_{radius}) can be different in X and Y directions:

I. Whole domain: CCM is applied on the whole domain by setting the search radii in X and Y directions to 5 pixels. To avoid the box from crossing the domain boundary, box dimensions are equal to X_{dim}-2X_{radius} and Y_{dim}-2Y_{radius}. For each velocity component (U_X and U_Y calculated separately), the velocity representative of the whole domain is obtained by finding the maximum correlated frame and calculating the displacement vector at the center of the box.

II. Smaller Boxes (block matching algorithm): the domain is split into boxes (e.g. 5 x 5 pixels) and for each box, (U_X, U_Y) are obtained by finding the maximum correlated frame and the displacement vector in (X, Y) direction at the center of each box. Smaller boxes result in a more granular velocity field in the domain.

The exhaustive (full) search method, which is the most computationally expensive block matching algorithm (Yang et al., 2012), is considered in this study. Alternatively, a three-step search (Li et al., 1994) can be used to overcome the cost of the exhaustive search. However, the

accuracy is limited by the reduced number of matching candidates. Note that the CCM is very sensitive to the box size and search radius. A small box size can result in insufficient pixels to distinguish the displacement maximum likelihood estimate. On the other hand, a large box size lead to insufficient number of vectors to describe the variation in cloud motion (Yang et al., 2012). Therefore, these parameters are chosen based on the expected range of the physical characteristics of the problem (typical wind speeds, etc.). Also, when there is more than one maximum correlated displacement (e.g. as is the case for periodic boundary conditions in the simulated dataset), the one with minimum displacement (equals to minimum velocity magnitude) is chosen.

3.2.2. Improving cloud speed resolution

Note that the resolution of the velocity vector that can be detected from two subsequent images in X and Y directions are $\Delta X/\Delta t$ and $\Delta Y/\Delta t$ respectively, where ΔX and ΔY are pixel spacing in X and Y directions and Δt is the time step. Therefore, valid values for the velocities are U_X = {$\Delta X/\Delta t$, 2 x $\Delta X/\Delta t$, …, X_{radius} x $\Delta X/\Delta t$}and U_Y = {$\Delta Y/\Delta t$, 2 x $\Delta Y/\Delta t$, …, Y_{radius} x $\Delta Y/\Delta t$}. Setting X_{radius} and Y_{radius} equal to 5 pixels ensure sufficient higher limit. However, if velocities are small such that, e.g. $U_X\,\Delta t < \Delta X$ or $U_Y\,\Delta t < \Delta Y$, the estimated velocity components U_X or U_Y will be equal to zero. Moreover, finer resolution (increment) of the velocity vectors are desirable. Computing velocities on additional images beyond two consecutive time steps improve the resolution and the lowest detectable cloud speed. In all cases and at each time step, the CCM results for horizons = 1, 2, 3, 4, and 5 (equal to Δt = 10, 20, 30, 40, and 50 s) are calculated and the one with the largest CCC is considered as the estimated cloud velocity. The cloud motion is considered to be zero if the averaged velocity vector equals zero up to horizon = 5.

3.2.3. QC

For local consistency as suggested by Hamill and Nehrkorn (1993), the vectors which differ from the local average (within a given distance) by more than a threshold are removed. Also, since the CCM for boxes with low variation, cloud fraction, and/or information content give unreliable and meaningless results, different QC parameters which were derived empirically are considered: (i) the variation ratio of each box (var_b) computed per Eq. 6 must be greater than 0.1, (ii) the box cloud cover fraction (the fraction of pixels in the box with $Kt_b < 0.85$) must be greater than 0.1. (iii) maximum CCC of the box max_{QC} must be greater than 0.8, (iv) cross-correlation QC ratio (r_{QC}) per Eq. 8 must be greater than 0.8. These QC parameters with the specified thresholds along with the local consistency are applied to the CCM for more accurate results (modified CCM).

3.2.4. Obtaining velocity field by objective analysis

After calculating the cloud velocity at the boxes in the domain and applying the QC, an objective analysis is applied to smooth the velocity field and produce a continuous cloud motion pattern. Therefore, each velocity vector is calculated using a multi-pass successive corrections scheme (3 passes) as suggested by Hamill and Nehrkorn (1993). The first estimate is

$$G1(i,j) = \frac{\sum_{g=1}^{nodes}[W^2 D(x_g, y_g)]}{\sum_{g=1}^{nodes} W}, \qquad \text{Eq. (9)}$$

where D is the displacement vector at location (x_g,y_g). W is the standard Cressman weighting function as

$$W_n = \begin{cases} \dfrac{R^2 - d_g^2}{R^2 + d_g^2} & \text{for } d_g < R \\ 0 & \text{for } d_g > R \end{cases}, \qquad \text{Eq. (10)}$$

where d_g is the distance between location (x_g,y_g) and pixel (i,j) and R is the radius of influence.

The second pass is calculated as

$$G2(i,j) = G1(i,j) + \frac{\sum_{g=1}^{nodes}\{W^2[D(x_g,y_g) - G1(x_g,y_g)]\}}{\sum_{g=1}^{nodes} W}. \qquad \text{Eq. (11)}$$

And the third pass (the estimated velocity field) is obtained in the same form as Eq. (11) with *G3* and *G2* replacing *G2* and *G1*, respectively. This 3 step scheme is applied for U_X and U_Y components separately.

4. Results

CCM and CSA methods are first validated on a moving 1D sinusoidal wave (not shown) and then using a 1D transect of the data in the along-wind and cross-wind directions (also not shown). Afterwards, the methods are applied to the 2D simulated data representing the simulated cumulus (RICO case) and stratocumulus cloud (CGILS case).

4.1. RICO simulation

The 1800 sec time series of all points in the domain (128x128 grid points) is analyzed. The reference cloud motion vector is obtained through the wind speeds at the average cloud height of 1031 m as 2.61 m/s at 180° (using the standard reference angle as the *X* axis and the positive angles calculated counter clockwise) in the whole domain. The reference cloud motion vector is used as a benchmark for the estimated cloud motion by applying the CCM and CSA. Fig. 1 shows *Kt* contours at selected time steps which illustrates cloud movement in the domain. While cloud motion is visually discernible to be along the negative *X* axis, motion detection is complicated by clouds changing shape and evaporating / forming in time.

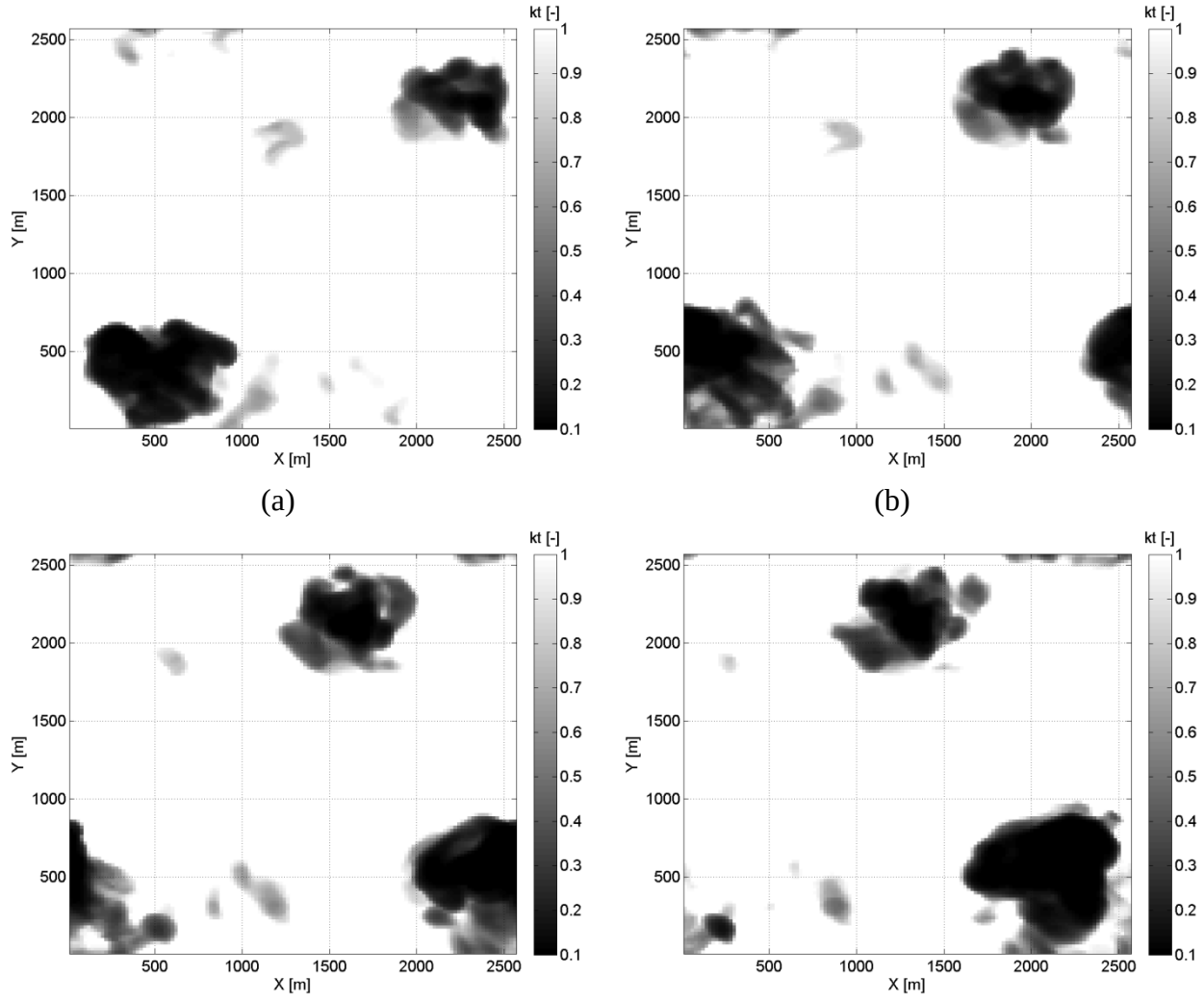


Fig. 1: *Kt* contour of the simulated dataset obtained from RICO simulation at different time steps; (a) 08:20:00, (b) 08:22:00, (c) 08:24:00, and (d) 08:26:00.

Fig. 2 shows time series of the domain averaged variation ratio and cloud cover fraction. As described in section 3.1.2, to ensure reliable cloud motion estimation, the points with low variation ratio or low cloud cover fraction are excluded from analysis.

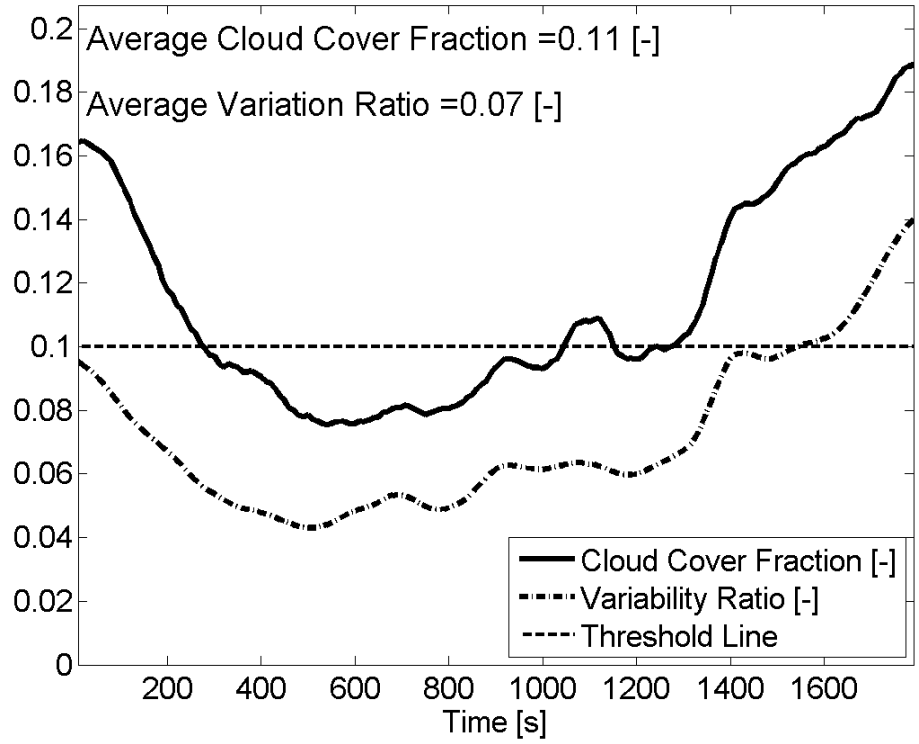


Fig. 2: Time series of the domain averaged cloud cover fraction (solid line) and variation ratio as calculated in Eq. 6 (dash-dot line). The text insert indicates the time average. The dashed line is the QC threshold that is applied to individual stations, not the domain-average shown here.

4.1.1. CSA results of the RICO simulation

The CSA is applied on the selected 16,384 sites (7.8×10^6 pairs after QC) and the 625 sites with highest variation ratio (5.8×10^4 pairs after QC) to estimate the cloud motion. Fig. 3 shows the time delay variation (Eq. 4) as a function of angle for $\max_{lag} = t_{dim}/6$ after application of the distance limit (Eq. 7). Fig. 3 confirms that choosing a subset of points (625 sites) is sufficient to obtain representative results.

The original CSA as presented in Section 3.1.1 (Inoue et al. (2012), not shown) using all the 16,384 sites yields a cloud motion of 1.34 m/s at 111° with error index e = 21%. The 33% improvement in speed and 52° absolute error reduction in cloud motion direction estimation as well as the reduction in the error index (from 21% to 8%) confirms that the proposed CSA is more accurate and robust than the original CSA and, therefore, the new robust CSA will be applied hereinafter.

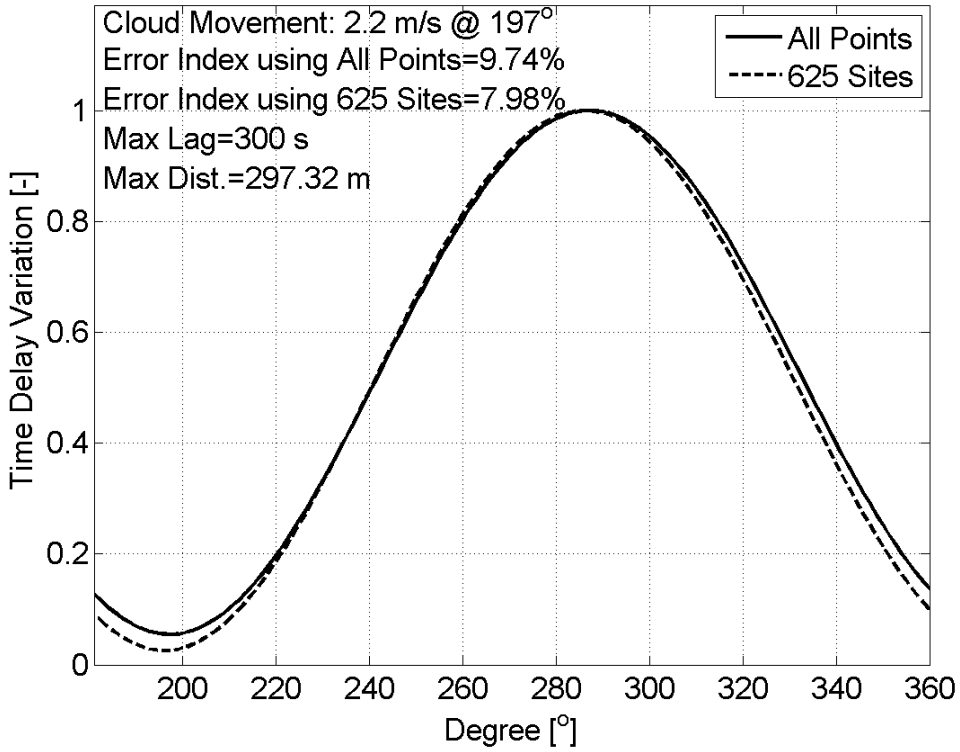


Fig. 3: Calculation of the time delay variation moment (Eq. 4) over all angles by considering pairs of sites within the distance limit using all points in the domain (16384 sites) and 625 sites with the highest variability. The caption indicates the calculated cloud direction, speed, and error index using all points as well as applied maximum time lag and maximum allowable distance between pairs of sites. The reference velocity at the average cloud height was 2.61 m/s at 180°.

Table 1 compares the results of cloud estimation of RICO simulation using the CSA by considering different sites. The results using all the points in the domain (16,384 points) are presented. Then, the QC steps described in Section 3.1.3 are applied one at a time and the accuracy of cloud motion vectors are quantified. Also, results using the sites with the highest variability (after QC) are presented. The CSA (after QC as described in Section 3.1.3) estimates the cloud motion in the whole domain with average relative mean bias error between the reference value and calculated velocity (rMBE) and difference between the reference cloud direction and calculated cloud motion direction (absolute error) equal to 15.7% and 17° respectively. All QC steps need to be applied to ensure accurate results. Also the average error index equals to 10% confirming the reliability of the method.

Table 1: Validation of cloud velocity for the RICO simulation by CSA. The effect of the QC steps described in Section 3.1.3 is quantified for the dataset with all 16,384 points: removing sites with low variability (step 1), removing sites with low variability and site pairs that are too distant (step 2), removing sites with low variability, site pairs that are too distant, and less correlated sites (step 3). 11,103 sites (68%) with low variation ratio or low cloud cover fraction, 5.1×10^6 (or 36%) of the remaining 1.4×10^7 pairs of sites (for the case with $max_{lag} = t_{dim}$), and 1.1×10^6 (or 12.5%) of the remaining 8.8×10^6 pairs of sites (for the case with $max_{lag} = t_{dim}$) are removed from the analysis in steps 1, 2, and 3, respectively. Also, results using only the 625 the sites with the highest variability (after QC) are presented. rMBE indicates the relative mean bias error between the reference value and CCM velocity and the absolute error indicates the difference between the reference cloud direction and calculated cloud direction. The average reference cloud motion in the whole domain is 2.61 m/s at 180°. The error index (*e*) quantifies the confidence of the results (Eq. 3).

Sites	Velocity			Direction	
	Speed (m/s)	rMBE (%)	*e* (%)	Angle (°)	Absolute Error (°)
All Points (16,384)	0.7	73.2	459.3	32	32
All Points, QC Step 1	1.2	54.0	284.1	201	21
All Points, QC Step 2	1.4	46.4	19.0	201	21
All Points, QC step 3	2.2	15.7	10.0	197	17
625 Sites	2.2	15.7	8.0	196	16

4.1.2. CCM results of the RICO simulation

The cloud velocity is estimated through the CCM applied to all grid points by considering the whole domain as a box. For more spatially granular estimation of the cloud motion, boxes with 5 x 5 pixels are chosen to calculate the velocity field using the objective analysis. QC criteria as described in Section 3.2.3 are applied to both the whole domain and the 5 x 5 boxes (modified CCM).

Fig. 4 shows the detailed modified CCM calculated cloud motion field at a single time step using boxes with 5 x 5 pixels through the whole domain. At this time step, the best result is achieved at horizon = 30 s. The cloud motion vector field is homogeneous.

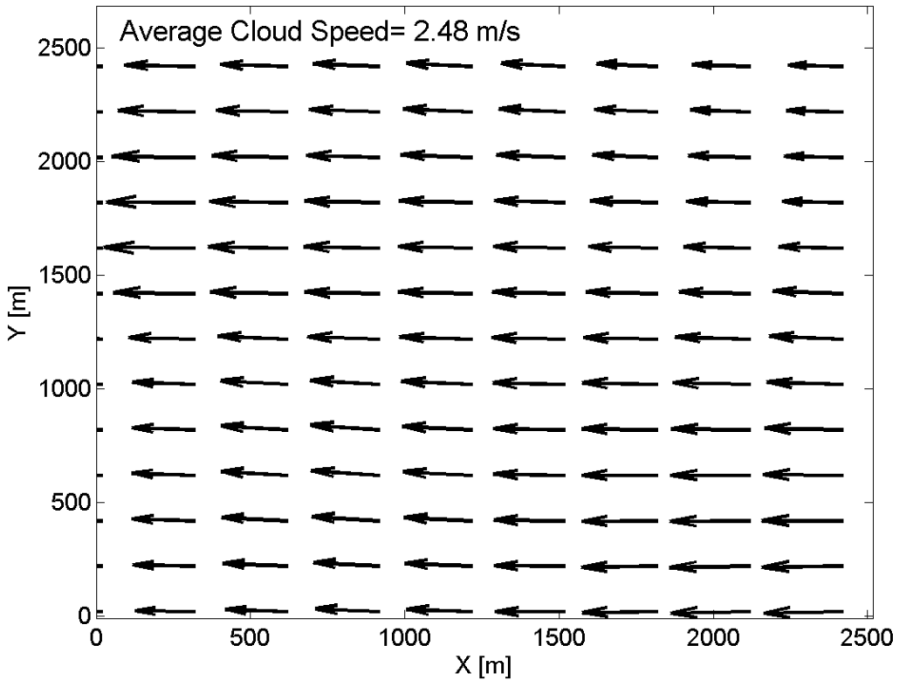
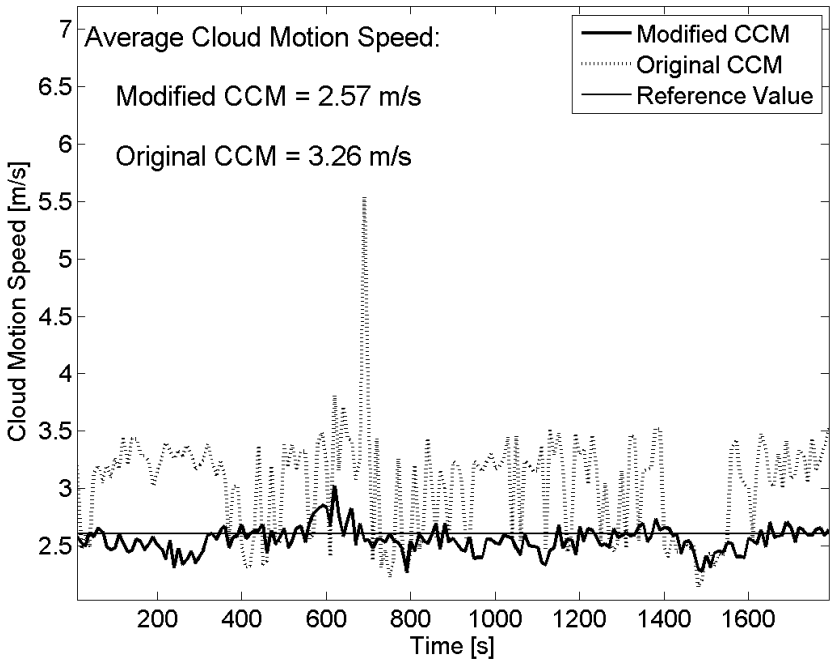
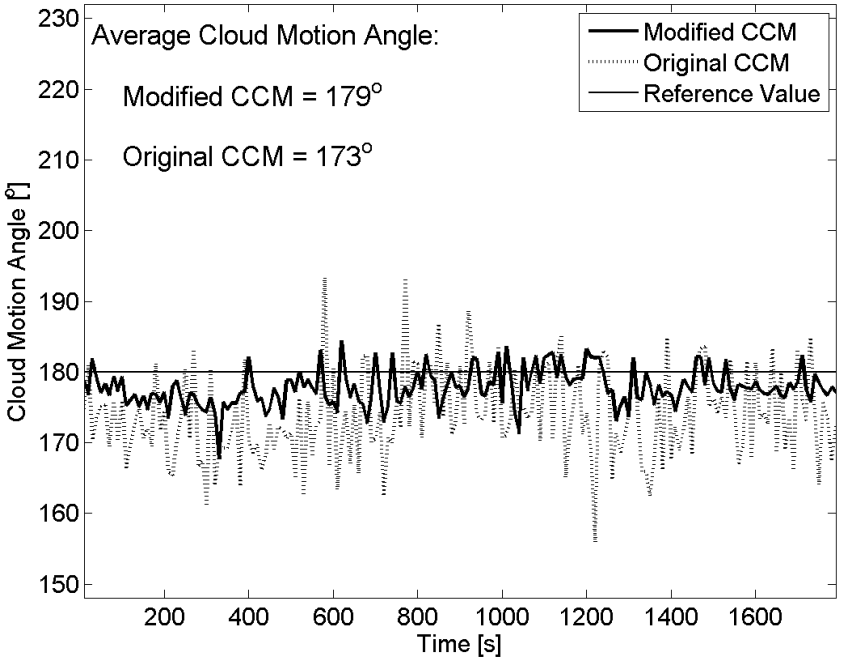


Fig. 4: Cloud velocity field calculation by modified CCM using boxes with 5 x 5 pixels through the whole domain at 08:27:40. The caption indicates the average calculated cloud speed.

Fig. 5 shows time series of the calculated cloud velocity by applying the local consistency and without QC criteria mentioned in Section 3.2.3 (original CCM) as well as after applying local consistency and QC (modified CCM). The improvements from applying QC are 21% in the speed and a 6° in the direction. Also the temporal variability in cloud motion vector is reduced. Therefore, the more robust and accurate modified CCM is applied hereinafter.



(a)



(b)

Fig. 5: Time series of CCM cloud (a) speed and (b) direction by using boxes with 5 x 5 pixels through the whole domain for original CCM (dashed line) and modified CCM (solid line). The dotted line shows the reference averaged wind (a) speed and (b) direction at average cloud height. The caption indicates the time-averaged cloud speed and direction.

To confirm consistency of the results and investigate the dependence on horizon, Table 2 shows the modified CCM results of the RICO simulation using the entire domain as a box as well as boxes with 5 x 5 pixels through the whole domain for horizons up to 50 s. In general, for this case with almost constant velocity over time, the CCM results using boxes with 5 x 5 pixels are in good agreement with actual cloud speed and direction with average rMBE and absolute error in the whole domain equal to 3.4% and 2° respectively. For horizons less than 30 s the cloud speed errors are larger, but speed and direction errors converge for horizons of 30 s or more.

The CCM that uses the entire domain as a box estimates actual cloud motion direction perfectly (180°). For cloud speed, however, the results at horizon = 1 are not that accurate (average rMBE equals to 23.4%). This confirms that, due to physical restriction, the CCM by choosing the entire domain as a box just gives a rough estimate and CCM method using boxes with 5 x 5 pixels through the domain should be applied to obtain more detailed results. Also, in such a case, considering CCM method at larger horizons provide more accurate results as suggested by Table 2, which illustrates the importance of calculating the cloud velocity using larger horizons. The optimum horizon is not generalizable, but depends on the cloud speed, time step, and station distance for a particular setup.

Table 2: CCM cloud velocity of RICO simulation using boxes with 5 x 5 pixels through the domain at different horizons *H*. rMBE indicates the relative mean bias error between the reference value and calculated velocity and the absolute error indicates the difference between the reference cloud direction and calculated cloud direction. The average reference cloud motion in the whole domain is 2.61 m/s at 180°.

Horizon	Resolution	Velocity		Direction	
		Speed (m/s)	rMBE (%)	Angle (°)	Absolute Error (°)
H = 1 (Δt = 10 s)	1 Box	2.00	23.4	180	0
	5x5 Boxes	2.42	7.3	178	2
H = 2 (Δt = 20 s)	1 Box	2.01	23.0	178	2
	5x5 Boxes	2.49	4.6	177	3
H = 3 (Δt = 30 s)	1 Box	2.18	16.5	179	1
	5x5 Boxes	2.55	2.3	178	2
H = 4 (Δt = 40 s)	1 Box	2.37	9.2	177	3
	5x5 Boxes	2.51	3.8	178	2
H = 5 (Δt = 40 s)	1 Box	2.33	10.7	179	1
	5x5 Boxes	2.63	-0.8	178	2

4.2. CGILS simulation

The CGILS case is split into the overcast and breakup sections of 10800 sec and 14400 sec duration with 10 sec resolution using all 96x96 equally spaced grid points in the domain (Figs. 6 & 7). While the reference cloud motion direction is constant at 280° over the CGILS simulation, the speed varies: during the overcast period the reference speed decreases from 9.9 to 7.5 m/s with an average of 8.6 m/s; during the breakup period the speed is more constant with 20 min means varying from 5.1 to 5.8 m/s with an average of 5.3 m/s.

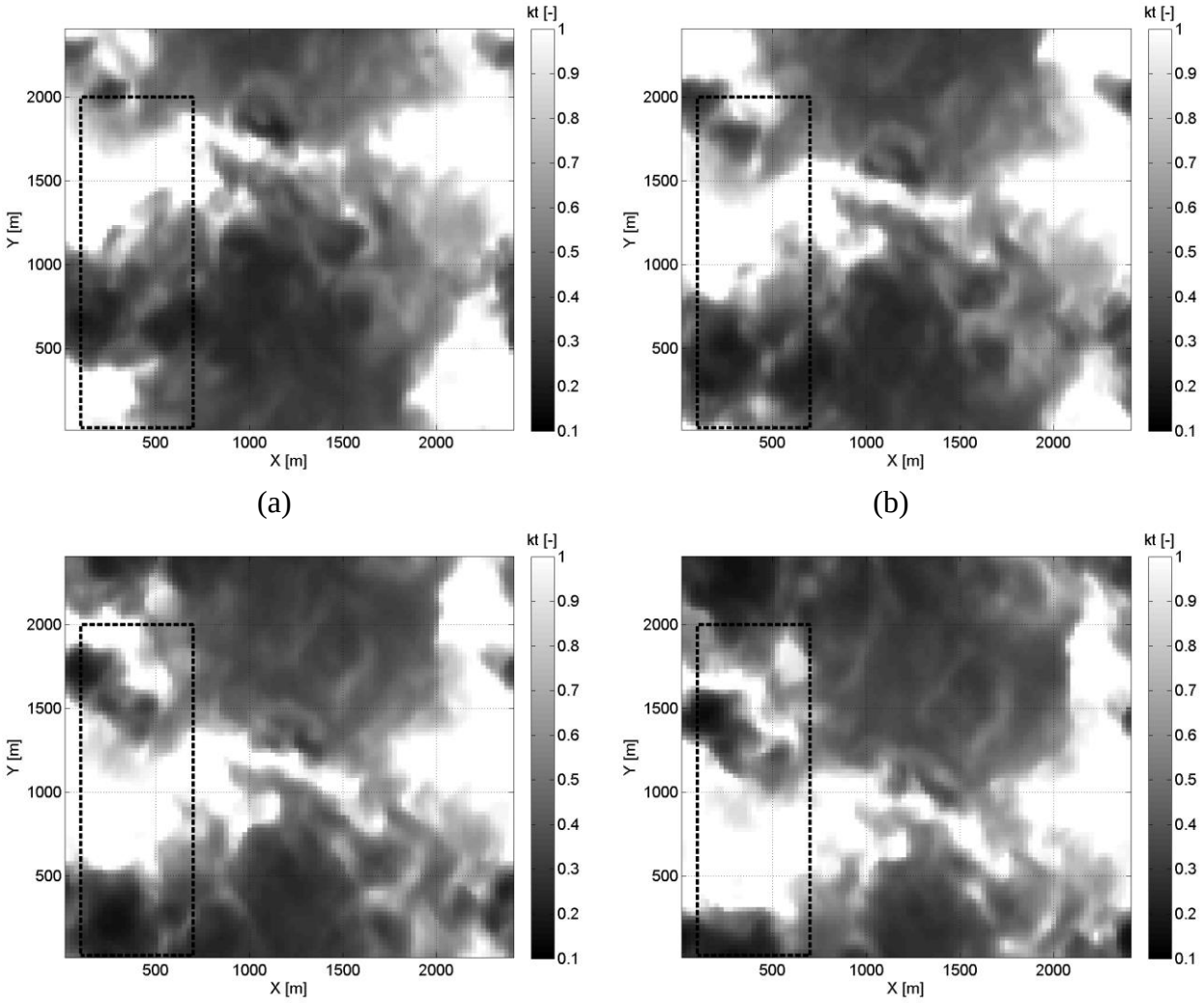


Fig. 6: *Kt* contour of the simulated dataset obtained from CGILS simulation at different time steps during the cloud breakup period; (a) 13:10:00, (b) 13:11:00, (c) 13:12:00, and (d) 13:13:00. The black dashed box represents the location of a subdomain which is used for further analysis in section 4.2.2.

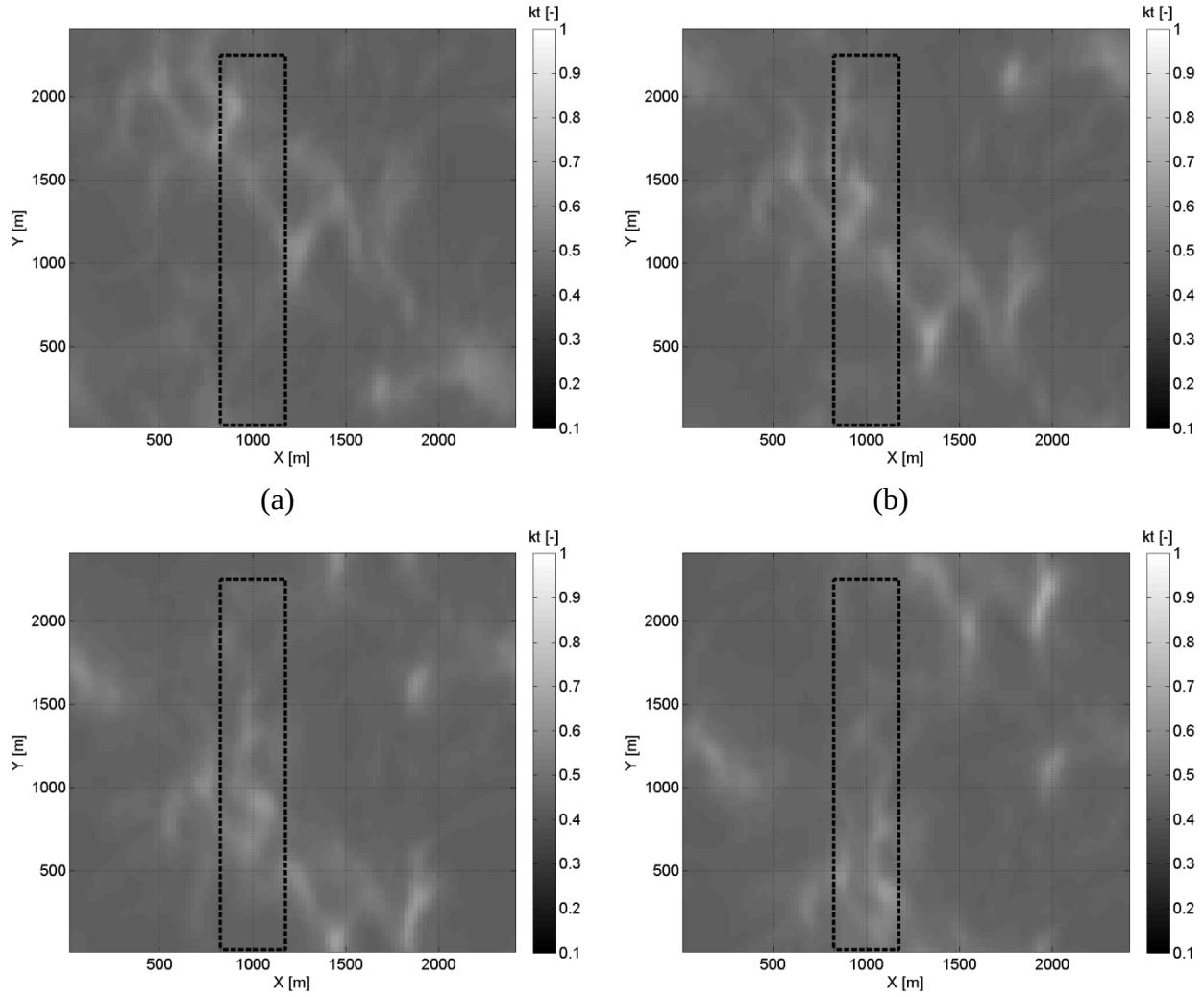
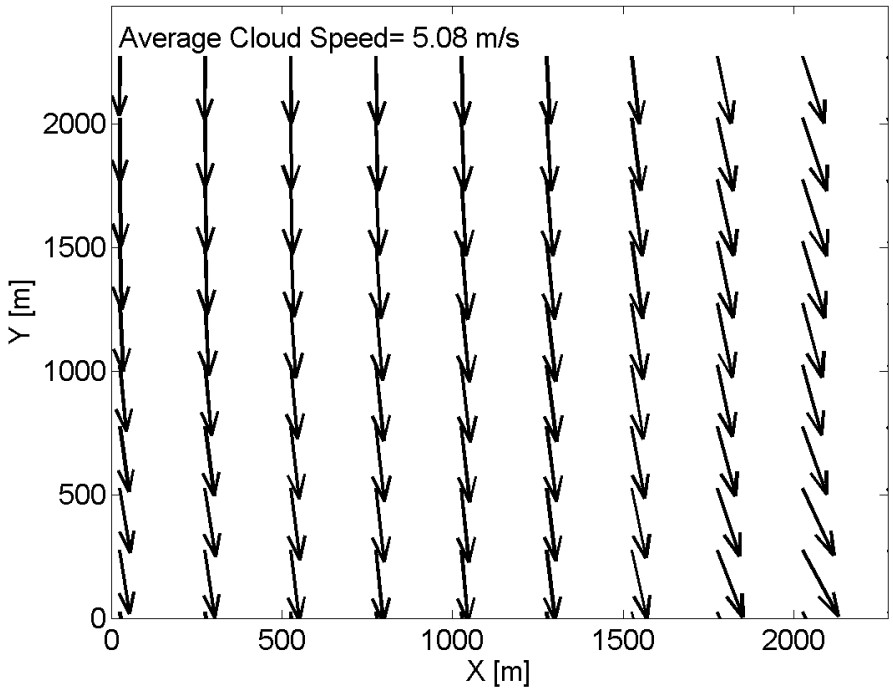


Fig. 7: Same as Fig. 6 but for CGILS simulation at different time steps during the overcast period; (a) 11:10:00, (b) 11:11:00, (c) 11:12:00, and (d) 11:13:00. The black dashed box represents the location of a subdomain which is used for further analysis.

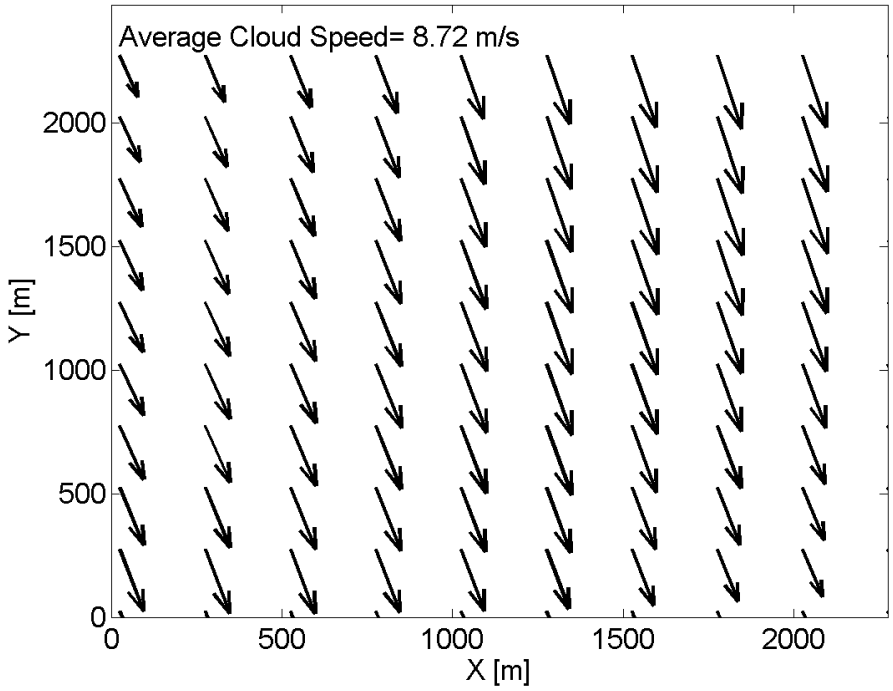
4.2.1. CSA and CCM results of the CGILS simulation

CCM and CSA are applied to estimate the cloud motion in CGILS simulation during the cloud breakup and overcast periods (Table 3). In CSA, 361 sites with highest variability within the whole domain are applied. CCM is applied by considering the entire domain as a box as well as using boxes with 5 x 5 pixels.

Similar to Fig. 5, Fig. 8 shows the detailed CCM calculated cloud motion field at a time step during each period using boxes with 5 x 5 pixels through the whole domain. The best cloud motion is estimated at horizon = 40 s and 10 s for the breakup (Fig. 8a) and overcast (Fig. 8b) periods respectively. While the CGILS cloud motion is not constant, the applied QC ensures local consistency and smoothness of the cloud motion field through the objective analysis.



(a)



(b)

Fig. 8: Same as Fig. 4 but at (a) 13:38:10 during the cloud breakup period and (b) 09:38:10 during the overcast period.

Table 3: Cloud velocity calculation of CGILS simulation during cloud breakup and overcast periods by CSA (similar to Table 1) and CCM (similar to Table 2). The averaged cloud cover fraction and variation ratio (Eq. 6) are, respectively, equal to 0.71 (0.67) and 0.37 (0.34) in the whole domain (subdomain) during the cloud breakup period while these values are equal to 1 (1) and 0.36 (0.16) in the whole domain (subdomain) during the overcast period respectively. Also, the average reference cloud motion in the whole domain is 5.32 m/s at 279° and 8.58 m/s at 281° during cloud breakup and overcast periods respectively.

Time Interval	Method	Domain	Resolution	Velocity			Direction	
				Speed (m/s)	rMBE (%)	*e* (%)	Angle (°)	Absolute Error (°)
1200 - 1500 h (breakup)	CSA	Whole Domain	361 Sites	0.29	94.5	82	334	55
		Subdomain	30 Sites	4.71	11.7	3	289	10
	CCM	Whole Domain	1 Box	5.12	3.8	-	270	9
			5x5 Boxes	5.01	5.8	-	276	3
		Subdomain	1 Box	5.18	2.6	-	275	4
			5x5 Boxes	5.26	1.1	-	276	3
0800 - 1200 h (overcast)	CSA	Whole Domain	361 Sites	1.50	82.5	19	353	72
		Subdomain	30 Sites	2.50	70.9	40	345	64
	CCM	Whole Domain	1 Box	8.69	-1.3	-	287	6
			5x5 Boxes	8.53	0.6	-	281	0
		Subdomain	1 Box	8.48	1.2	-	283	2
			5x5 Boxes	8.61	-0.3	-	279	2

Table 3 confirms that, during the cloud breakup (with average cloud cover fraction equal to 0.71 in the whole domain), CCM results are in agreement with the average reference cloud speed (5.32 m/s) and direction (279°), with average rMBE of speed and absolute error of direction equal to 4.8% and 6° respectively. On the other hand, the CSA results are not accurate during cloud breakup; the average rMBE and absolute error are equal to 94.5% and 55° for the cloud motion speed and direction respectively. However, the average error index of 82% confirms that the results are not reliable. In the overcast period (with cloud cover fraction equal to 1), CCM results are accurate with average rMBE and absolute error equal to -0.3% and 3° respectively. In this period, the CSA results are again not accurate; the average rMBE and absolute error equal to 82.5% and 72° for the cloud motion speed and direction respectively and the average error index equals to 19%.

4.2.2. CCM results of a subdomain of the CGILS simulation

The quality control ensures reliable results as the points with low information content (including low variability and low cloud cover fraction) are removed from the analysis. Given the spatial homogeneity in the domain and boundary conditions, the cloud velocity field should be consistent in the whole domain. To evaluate the performance of the CCM in a smaller region and with less available data, subdomains with 25 x 80 (the black box in Fig. 6) and 15 x 90 grid points (the black box in Fig. 7) are considered during cloud breakup and overcast periods respectively. While the locations and sizes of the subdomains were chosen arbitrarily they were aligned with the cloud motion direction to resemble an idealized 1-dimensional setup.

Boxes with 5 x 5 pixels as well as a box covering the entire subdomain are considered in CCM. Similar to the analysis for the whole domain, the cloud fields in the subdomains are

estimated using CCM up to horizon = 50 sec and the best CCM results are presented in Table 3. The cloud motion speed and direction is estimated using CCM by applying 5 x 5 boxes in the subdomains with average rMBE (absolute error) equal to 5.8% (3°) and -0.3% (2°) during the cloud breakup and overcast periods respectively. The results confirm the accuracy of the CCM method in the subdomain and the consistency with the results for the whole domain.

4.2.3. CSA results of a subdomain of the CGILS simulation

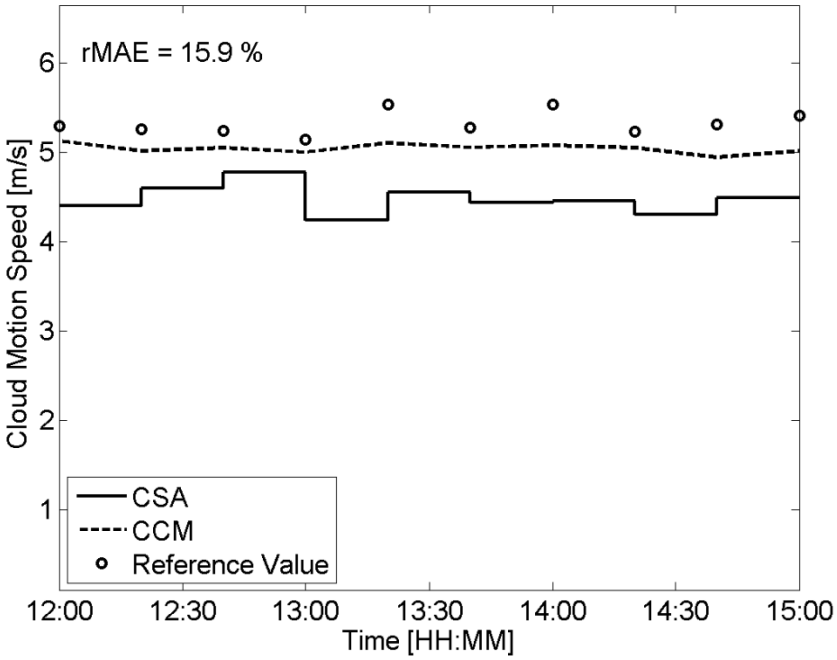
CSA results are not accurate in the whole domain and, therefore, it should be tested if the method performs better in the subdomains. The subdomains described in Section 4.2.2 are considered for the CGILS simulation during the cloud breakup and overcast periods. 30 sites with highest variability through the subdomains are applied in CSA.

Table 3 shows the cloud estimation results for the subdomains using CSA. During the cloud breakup, in comparison with the results for the whole domain, the CSA method provides more accurate estimates of the cloud motion in the subdomain with average rMBE and absolute error equal to 11.7% and 10° for the cloud motion speed and direction respectively and average error index equals to 3%. In the overcast period, similar to the results for the whole domain, the CSA results are not accurate with the average rMBE and absolute error equal to 70.9% and 64° for the cloud motion speed and direction respectively and the average error index equals to 40%.

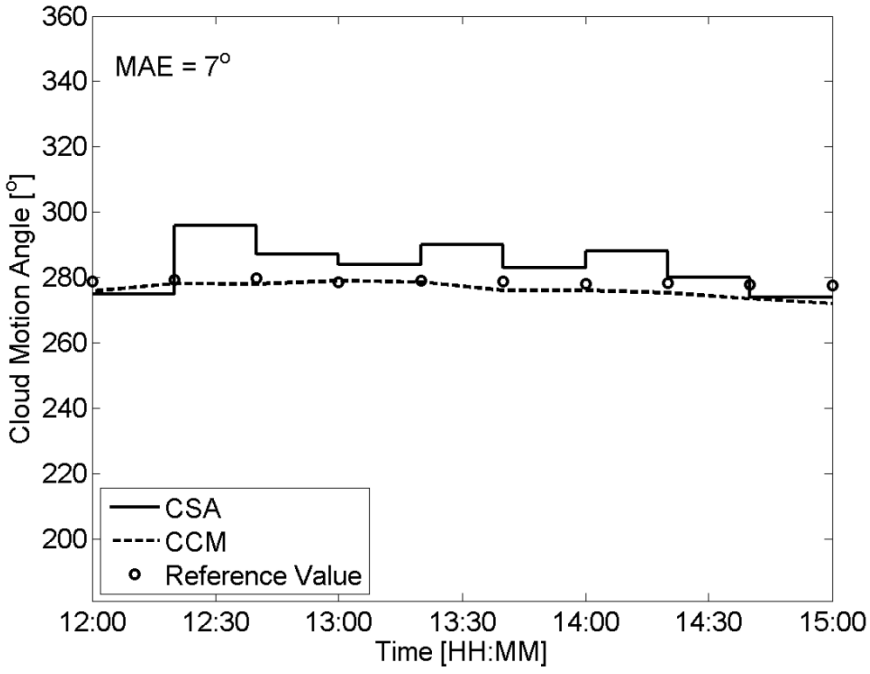
4.2.4. Analysis of subintervals of the CGILS simulation

The results of the CGILS simulation during the cloud breakup show that CSA is unable to detect the cloud motion pattern in the whole domain while this pattern is detected in the smaller subdomain. Moreover, during the overcast period with cloud cover fraction equals to 1, results illustrate that CSA is unable to detect the cloud motion pattern, even in the smaller subdomain. This is related to the fact that, in CSA, a constant velocity vector (as the average of the velocity field over the whole period under investigation) is assumed and temporal changes in cloud speed and/or direction degrade the accuracy of the CSA method. For high cloud cover fractions, the probability of the temporal changes in cloud motion increases and CSA is unable to detect the cloud motion pattern in these situations.

To understand the maximum length of subintervals for the CSA, the entire time series in each period (during cloud breakup and overcast) is split into subintervals and CSA is applied to estimate the cloud motion in each subinterval. The lengths of subintervals are varied from 5 to 120 min in 5 min increments. Figs. 9 & 10 show the estimated cloud motion speed and direction using CSA in 20 min subintervals. Also, average relative mean absolute error (rMAE) between the reference cloud speed and the estimated cloud speed using CSA at each subinterval and mean absolute error (MAE) between the reference cloud direction and cloud direction estimated by CSA at each subinterval are calculated. Using the 20 min subintervals, the CSA estimates the cloud motion in the cloud breakup (overcast) period with average rMAE and rMBE of speed equal to 15.8% (22.2%) and 15.8% (20.7%) respectively as well as MAE and absolute error of direction equal to 7° (23°) and 5° (23°) respectively. While this is a large improvement over using CSA for the entire time series; the CCM results are still more accurate in both cases.

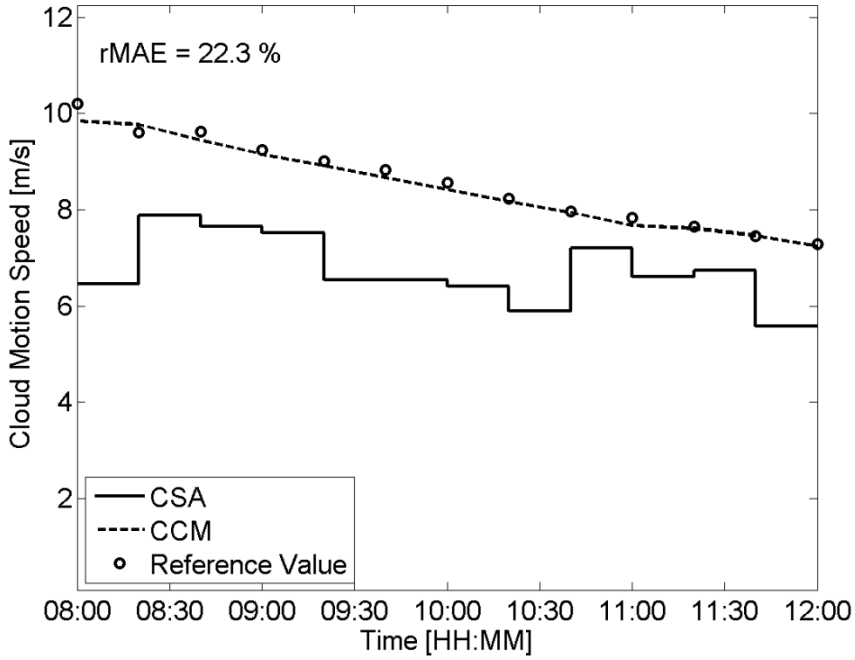


(a)

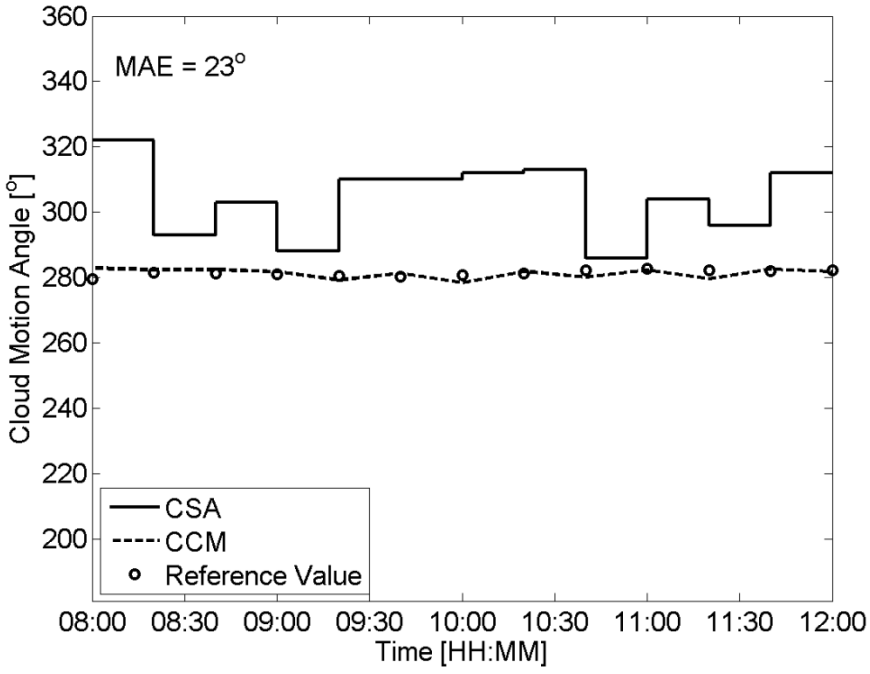


(b)

Fig. 9: Time series of cloud (a) speed and (b) direction by CSA using 20 min subintervals during the cloud breakup period through the whole domain. The CCM results and reference values are illustrated as well. The caption indicates the average relative mean absolute error (rMAE) between cloud speed calculated by CSA and the reference value (in a) and the mean absolute error (MAE) between cloud direction calculated by CSA and the reference cloud direction.



(a)



(b)

Fig. 10: Same as Fig. 9 but for the overcast period.

Fig. 11 shows the cloud motion estimation results for CGILS simulation during the cloud breakup and overcast conditions. For each subinterval length, the cloud motion speed and direction are calculated in each subinterval and the averaged values over all the subintervals (entire timeseries) are presented. The analysis is performed in the whole domains as well as the

subdomains in each period. The results confirm that during the cloud breakup period, the cloud motion speed can be estimated with average rMBE less than 20% by choosing subintervals up to 120 min. However, during the overcast period, choosing subintervals larger than 30 min will lead to inaccurate (rMBE > 27%) and unreliable results in both the whole domain and the subdomain.

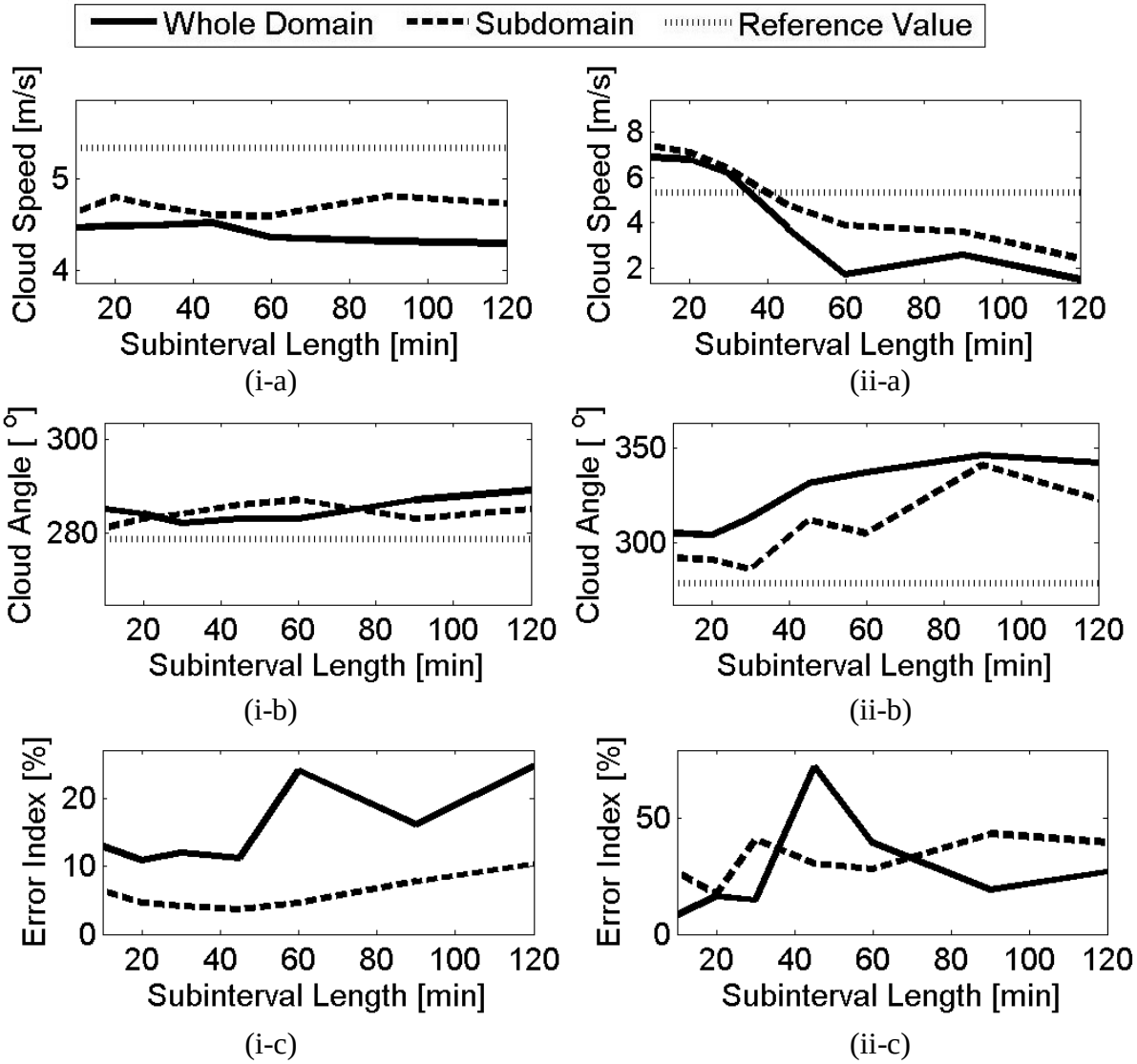


Fig. 11: Cloud velocity estimation of CGILS simulation during cloud (i) breakup and (ii) overcast periods using CSA at different subinterval lengths for the whole domain and subdomain; (a) cloud speed, (b) cloud direction, and (c) error index. For each case, the values are averaged over all the subintervals and the reference values are illustrated as well.

5. Conclusions

The cloud motion has been estimated by CSA and CCM using two spatially and temporally resolved simulated irradiance datasets generated from large eddy simulation (LES). CCM estimates cloud motion by comparing correlation between *Kt* data at two or more time steps. CCM is applied by considering the whole domain as a box. For spatially resolved estimation of

the cloud motion the domain is subdivided into boxes with 5 x 5 pixels and the velocity field is post-processed using objective analysis. Quality control (QC) is performed to remove data with low information content (including low variability and low cloud cover fraction), which ensures reliable results. Moreover, estimation of the cloud cover fraction and variation ratio may help to improve cloud motion estimation which, in turn, will increase the accuracy of solar forecasting algorithms.

In general, CCM results are accurate in all cases and the CCM method using boxes with 5 x 5 pixels in the domain provides more reliable results for the cloud motion speed. In this paper, CCM results are calculated up to a horizon = 50 sec, which was empirically found to be sufficient to obtain the best results. However, in general, the best forecast horizon is mainly a function of cloud speed and distance between the sites. In real situations with unknown cloud speed, the maximum horizon can be adjusted or even detected automatically according to the specifications of the problem. Specifically the variation ratio and cloud cover fraction, where the former depends on how frequently the clouds are passing while the latter represents the fraction of the area shaded by clouds in the domain, affect the optimum horizon. As a future work, by analyzing these variables for different datasets and in different situations, a general relationship between the maximum horizon and these variables may be obtained.

CSA estimates the cloud speed and direction by cross-spectral analysis of *Kt* data at all grid points in the domain. Computational costs can be reduced by choosing sites with the highest local variability (625 and 361 sites in the whole domain of RICO and CGILS, respectively) and it was shown that this does not negatively impact the accuracy of the results. Since there are usually less sites available in reality, the number of sites was further reduced (not shown) to as few as 36 (25) sites for RICO (CGILS) and the accuracy of the results was maintained. CSA results are reliable for the cases with low cloud cover fractions; for the RICO case (for both the whole domain and the subdomain) and for the subdomain in the CGILS simulation during the cloud breakup. Since the entire time series is considered simultaneously in the CSA approach, a constant velocity vector (as the average of the velocity field over the whole period under investigation) is assumed, i.e., the cloud motion is assumed to be spatially homogeneous and steady or at most slowly varying in time, such that it could be assumed to be piecewise steady. Therefore, temporal changes in cloud speed and/or direction degrade the accuracy of the CSA method. Moreover, in the cases with relatively high cloud cover fractions, the problem becomes more complicated and the probability of temporal changes in cloud motion increases. Therefore, CSA is unable to detect the cloud motion pattern for the whole time series in these situations, however, CSA estimates cloud motion more accurately in shorter time intervals (in which the cloud motion can be assumed to be spatially homogeneous and steady). To obtain acceptable results in CSA, the maximum length of the subintervals are observed to be 120 and 30 min for CGILS simulation during the cloud breakup and overcast periods respectively.

For generalization of CSA application, it would be helpful to detect the maximum length of time series or allowable variation in cloud motion for which acceptable results are achieved. This can be performed by considering some parameters including the domain size as well as the spatial and temporal resolutions of the *Kt* data. Therefore, for a given dataset, the maximum allowable length of time series can be quantified by estimating a threshold for cloud cover fraction and variation ratio of *Kt* time series. Datasets with different resolutions in different weather conditions are required to obtain a general relation between these parameters.

Although the results shown here are obtained using the simulated data where the cloud motion is either constant (RICO) or slightly decreasing (CGILS), the proposed modified CSA and CCM

can be applied for a wide range of cloud motion direction and/ or speed as shown by Jamaly and Kleissl (2017). They applied the anisotropic kriging method for spatio-temporal interpolation and irradiance forecast for the RICO simulation as well as the CGILS simulation during the overcast and cloud breakup periods. They also applied anisotropic kriging to ground measured irradiance and power output data in California where the cloud motion had to be estimated during the process. Accurate estimation of cloud motion by the modified CSA and CCM resulted in average normalized root mean squared irradiance errors (nRMSE) of 7.92% representing a 66% relative improvement over irradiance persistence. The accuracy of the anisotropic kriging method is sensitive to accurate estimation of the cloud motion and, therefore, obtaining large improvement over the persistence model for real data and for different weather conditions confirms the accuracy and reliability of kriging in general, and the modified CCM and CSA in specific.

References

- Arias-Castro, E., Kleissl, J., and Lave, M., (2014), "A Poisson model for anisotropic solar ramp rate correlations", Solar Energy, 101:192-202. DOI: 10.1016/j.solener.2013.12.028.
- Arking, A., Lo, R.C., and Rosenfeld, A. (1978), "A Fourier approach to cloud motion estimation", Journal of Applied Meteorology, 17(6), 735-744. DOI: 10.1175/1520-0450(1978)017<0735:AFATCM>2.0.CO;2.
- Bedka, K.M., and Mecikalski, J.R., (2005), "Application of satellite-derived atmospheric motion vectors for estimating mesoscale flows", Journal of Applied Meteorology, 44 (11), 1761-1772. DOI: 10.1175/JAM2264.1.
- Bosch, J.L., and Kleissl, J., (2013), "Cloud motion vectors from a network of ground sensors in a solar power plant", Solar Energy 95, 13–20. DOI: 10.1016/j.solener.2013.05.027.
- Bosch, J.L., Zheng, Y., and Kleissl, J., (2013), "Deriving cloud velocity from an array of solar radiation measurements", Solar Energy 87, 196–203. DOI: 10.1016/j.solener.2012.10.020.
- Chow, C.W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., and Washom, B., (2011), "Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed", Solar Energy 85, 2881–2893. DOI: 10.1016/j.solener.2011.08.025.
- Chow, C.W., Belongie, S., and Kleissl, J., (2015), "Cloud motion and stability estimation for intra-hour solar forecasting", Solar Energy 115, 645-655. DOI: 10.1016/j.solener.2015.03.030.
- Engerer, N.A., and Mills, F.P., (2014), "K PV: A clear-sky index for photovoltaics", Solar energy, 2014 Jul 31;105:679-93. DOI: 10.1016/j.solener.2014.04.019.
- Escrig, H., Batlles, F.J., Alonso, J., Baena, F.M., Bosch, J.L., Salbidegoitia, I.B., and Burgaleta, J.I., (2013), "Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast", Energy, 55, 853-859. DOI: 10.1016/j.energy.2013.01.054.
- Farnebäck, G., (2003), "Two-frame motion estimation based on polynomial expansion" in: Bigun, J., and Gustavsson, T., Image Analysis. SCIA 2003. Lecture Notes in Computer Science, vol. 2749. Springer, Berlin, Heidelberg. DOI: 10.1007/3-540-45103-X_50.
- Fuh, C.S., and Maragos, P. (1991), "Motion displacement estimation using an affine model for image matching", Optical Engineering, 30(7), 881-887. DOI: 10.1117/12.55885.
- Fung, V., Bosch, J.L., Roberts, S.W., and Kleissl, J., (2013), "Cloud speed sensor", Atmos. Meas. Tech. Discuss. 6 (5), 9037–9059.
- Ghonima, M.S., Heus, T., Norris, J.R., and Kleissl, J., (2016), "Factors Controlling Stratocumulus Cloud Lifetime over Coastal Land", J. Atmospheric Sciences (73), 2961-2983. DOI: 10.1175/JAS-D-15-0228.1.
- Hamill, T.M., and Nehrkorn, T., (1993), "A short-term cloud forecast scheme using cross correlations", Weather & Forecasting, 8-4, pp. 401-411. DOI: 10.1175/1520-0434(1993)008<0401:ASTCFS>2.0.CO;2.
- Hammer, A., Heinemann, D., Lorenz, E., and Lückehe, B., (1999), "Short-term forecasting of solar radiation: a statistical approach using satellite data", Solar Energy 67, 139–150. DOI: 10.1016/S0038-092X(00)00038-4.
- Hess, M., Koepke, P., and Schult, I., (1998), "Optical properties of aerosols and clouds", *Bull. Amer. Meteor. Soc*., 79, 831-44. DOI: 10.1175/1520-0477(1998)079<0831:OPOAAC>2.0.CO;2.

- Hoff, T.E., and Perez, R., (2010), "Quantifying PV power output variability", Solar Energy 84, 1782–1793. DOI: 10.1016/j.solener.2010.07.003.
- Huang, H., Xu, J., Peng, Z., Yoo, S., Yu, D., Huang, D., and Qin, H., (2013), "Cloud motion estimation for short term solar irradiation prediction", Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference, pp. 696–701. DOI: 10.1109/SmartGridComm.2013.6688040.
- Inoue, T., Sasaki, T., and Washio, T., (2012), "Spatio-temporal kriging of solar radiation incorporating direction and speed of cloud movement", 26th Annual Conference of the Japanese Society for Artificial Intelligence, Yamaguchi city.
- Jamaly, M., and Kleissl, J., (2017), "Spatiotemporal Interpolation and Forecast of Irradiance Data Using Kriging", Solar Energy, Vol. 158, pp. 407-423. DOI: 10.1016/j.solener.2017.09.057.
- Lave, M., and Kleissl, J., (2013), "Cloud Speed Impact on Solar Variability Scaling - Application to the Wavelet Variability Model", Solar Energy, 91:11-21. DOI: 10.1016/j.solener.2013.01.023.
- Leese, J.A., Novak, C.S., and Clark, B.B., (1971), "An automated technique for obtaining cloud motion from geosynchronous satellite data using cross correlation", Journal of Applied Meteorology 10, 118–132. DOI: 10.1175/1520-0450(1971)010<0118:AATFOC>2.0.CO;2.
- Li, R., Zeng, B., and Liou, M.L., (1994), "A New Three-Step Search Algorithm for Block Motion Estimation", IEEE Transactions on Circuits and Systems for Video Technology, VOL. 4, NO. 4, 438-442. DOI: 10.1109/76.313138.
- Lorenz, E., Hammer, A., and Heinemann, D., (2004), "Short term forecasting of solar radiation based on satellite data", Proc. ISES Europe Solar Congress EUROSUN2004, Freiburg, Germany.
- Lorenzo, A.T., Holmgren, W.F., Leuthold, M., Kim, C.K., Cronin, A.D., and Betterton, E.A., (2014), "Short-Term PV Power Forecast Based on a Real-Time Irradiance Monitoring Network", IEEE Photovoltaic Specialist Conference (PVSC), DOI: 10.1109/PVSC.2014.6925212. DOI: 10.1109/PVSC.2014.6925212.
- Marcos, J., Marroyo, L., Lorenzo, E., Alvira, D., and Izco, E., (2011), "Power output fluctuations in large scale PV plants: one year observations with one second resolution and a derived analytic model", Progress in Photovoltaics: Research and Applications, 19-2: 218-227. DOI: 10.1002/pip.1016.
- Marquez, R., and Coimbra, C.F.M., (2013), "Intra-hour DNI forecasting based on cloud tracking image analysis", Solar Energy, 91, 327–336. DOI: 10.1016/j.solener.2012.09.018.
- Marquez, R., Pedro, H.T.C., and Coimbra, C.F.M., (2013), "Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs", Solar Energy 92, 176–188. DOI: 10.1016/j.solener.2013.02.023.
- Menzel, W.P., (2001), "Cloud tracking with satellite imagery: from the pioneering work of Ted Fujita to the present", Bull. Am. Meteorol. Soc. 82 (1), 33–47. DOI: 10.1175/1520-0477(2001)082<0033:CTWSIF>2.3.CO;2.
- Perez, R., and Hoff, T.E., (2013). Chapter 10 – SolarAnywhere forecasting. In: Kleissl, J. (Ed.), "Solar Energy Forecasting and Resource Assessment", Academic Press, Boston, pp. 233–265.
- Perez, R., Kivalov, S., Schlemmer, J., Hemker Jr., K., Renn´e, D., and Hoff, T.E., (2010), "Validation of short and medium term operational solar radiation forecasts in the US", Solar Energy 84, 2161–2172. DOI: 10.1016/j.solener.2010.08.014.
- Quesada-Ruiz, S., Chu, Y., Tovar-Pescador, J., Pedro, H.T.C., and Coimbra, C.F.M., (2014), "Cloud-tracking methodology for intra-hour DNI forecasting", Sol. Energy 102, 267–275. DOI: 10.1016/j.solener.2014.01.030.
- Shinozaki, K., Yamakawa, N., Sasaki, and T., Inoue, T., (2014), "Areal solar irradiance estimated by sparsely distributed observations of solar irradiance", IEEE Transactions on Power Systems, TPWRS-00139-2014.R1. DOI: 10.1109/TPWRS.2015.2393636.
- Slingo, A., (1989), "A GCM Parameterization of the shortwave radiative properties of water clouds", American Meteorological Society, vol. 46, No. 10, pp. 1419-1427. DOI: 10.1175/1520-0469(1989)046<1419:AGPFTS>2.0.CO;2.
- vanZanten, M.C., Stevens, B., Nuijens, L., Siebesma, A.P., Ackerman, A.S., Burnet, F., Cheng, A., Couvreux, F., Jiang, H., Khairoutdinov, M., Kogan, Y., Lewellen, D.C. , Mechem, D., Nakamura, K., Noda, A., Shipway, B.J., Slawinska, J., Wang, S., and Wyszogrodzki, A., (2011), " Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO", J. Adv. Model. Earth Syst., Vol. 3, Art. M06001. DOI: 10.1029/2011MS000056.

- Vega-Riveros, J.F., and Jabbour, K. (1989), "Review of motion analysis techniques", IEE Proceedings I (Communications, Speech and Vision), 136(6), 397-404. DOI: 10.1049/ip-i-2.1989.0060.
- Yang, H., Kurtz, B., Nguyen, D., Urquhart, B., Chow, C.W., Ghonima, M., and Kleissl, J., (2014a), "Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego", Solar Energy 103, 502-524. DOI: 10.1016/j.solener.2014.02.044.
- Yang, D., Dong, Z., Reindl, T., Jirutitijaroen, P., and Walsh, W.M., (2014b), "Solar irradiance forecasting using spatio–temporal empirical Kriging and vector autoregressive models with parameter shrinkage", Solar Energy, 103(0):550–562. DOI: 10.1016/j.solener.2014.01.024.
- Yang, D., Walsh, W.M., Zibo, D., Jirutitijaroen, P., and Reindl, T.G., (2012), "Block Matching Algorithms: Their Applications and Limitations in Solar Irradiance Forecasting", Energy Procedia 33 (2013) 335 – 342. DOI: 10.1016/j.egypro.2013.05.074.
- Zhang, M., Bretherton, B., Lock, A., Webb, M., Khairoutdinov, M., and Cheng, A., (2009), "Class Specification for the CFMIP-GCSS Study of Cloud Feedback Mechanisms by Using SCM/CRM/LES Models", on Mar. 25, 2009 at: http://atmgcm.msrc.sunysb.edu/cfmip_figs/Case_specification.html.